

RECEIVED

MAR -7 2002

TECH CENTER 1600/2900



1652

RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/727,769A

TIME: 11:27:23

Input Set : A:\pto.vsk.txt

Output Set: N:\CRF3\02212002\I727769A.raw

ENTERED

3 <110> APPLICANT: Amano Enzyme, Inc.

5 <120> TITLE OF INVENTION: NOVEL PROTEIN-DEAMIDATING ENZYME, MICROORGANISM PRODUCING THE SAME, GENE

6 ENCODING THE SAME, PRODUCTION PROCESS THEREFOR, AND USE THEREOF

8 <130> FILE REFERENCE: Q62106

10 <140> CURRENT APPLICATION NUMBER: 09/727,769A

11 <141> CURRENT FILING DATE: 2000-12-04

13 <150> PRIOR APPLICATION NUMBER: JP 11-345044

14 <151> PRIOR FILING DATE: 1999-12-03

16 <160> NUMBER OF SEQ ID NOS: 11

18 <170> SOFTWARE: PatentIn version 3.1

20 <210> SEQ ID NO: 1

21 <211> LENGTH: 20

22 <212> TYPE: PRT

23 <213> ORGANISM: Cryseobacterium sp. No. 9670

25 <400> SEQUENCE: 1

27 Leu Ala Ser Val Ile Pro Asp Val Ala Thr Leu Asn Ser Leu Phe Asn

28 1 5 10 15

31 Gln Ile Lys Asn

32 20

35 <210> SEQ ID NO: 2

36 <211> LENGTH: 20

37 <212> TYPE: PRT

38 <213> ORGANISM: Cryseobacterium sp. No. 9670

40 <400> SEQUENCE: 2

42 Ser Pro Ser Asn Ser Tyr Leu Tyr Asp Asn Asn Leu Ile Asn Thr Asn

43 1 5 10 15

46 Cys Val Leu Thr

47 20

50 <210> SEQ ID NO: 3

51 <211> LENGTH: 20

52 <212> TYPE: DNA

53 <213> ORGANISM: Artificial Sequence

55 <220> FEATURE:

56 <223> OTHER INFORMATION: sense primer

58 <220> FEATURE:

59 <221> NAME/KEY: misc_feature

60 <222> LOCATION: (3)..(3)

61 <223> OTHER INFORMATION: n = inosine

64 <220> FEATURE:

65 <221> NAME/KEY: misc_feature

66 <222> LOCATION: (6)..(6)

67 <223> OTHER INFORMATION: n = inosine

70 <220> FEATURE:

RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/727,769A

TIME: 11:27:23

Input Set : A:\pto.vsk.txt

Output Set: N:\CRF3\02212002\I727769A.raw

71 <221> NAME/KEY: misc_feature
 72 <222> LOCATION: (9)..(9)
 73 <223> OTHER INFORMATION: n = inosine
 76 <220> FEATURE:
 77 <221> NAME/KEY: misc_feature
 78 <222> LOCATION: (15)..(15)
 79 <223> OTHER INFORMATION: n = a, c, g or t
 82 <400> SEQUENCE: 3
 83 gchwsngtna thccngaygt 20
 86 <210> SEQ ID NO: 4
 87 <211> LENGTH: 20
 88 <212> TYPE: DNA
 89 <213> ORGANISM: Artificial Sequence
 91 <220> FEATURE:
 92 <223> OTHER INFORMATION: anti-sense primer
 94 <220> FEATURE:
 95 <221> NAME/KEY: misc_feature
 96 <222> LOCATION: (3)..(3)
 97 <223> OTHER INFORMATION: n = a, c, g or t
 100 <220> FEATURE:
 101 <221> NAME/KEY: misc_feature
 102 <222> LOCATION: (12)..(12)
 103 <223> OTHER INFORMATION: n = a, c, g or t
 106 <400> SEQUENCE: 4
 107 arnacrcart tngtrtttdat 20
 110 <210> SEQ ID NO: 5
 111 <211> LENGTH: 555
 112 <212> TYPE: DNA
 113 <213> ORGANISM: Cryseobacterium sp. No. 9670
 115 <400> SEQUENCE: 5
 116 ttggcgagtg taattcctga ttagctaca ttaaattctt tattcaatca aataaagaat 60
 118 cagtcttgcg gtacctctac ggcgtcctca ccatgcatca cattcagata tcctgtagac 120
 120 ggatgttatg caagagccca taagatgaga caaatcttaa tgaacaacgg ctatgactgt 180
 122 gaaaaacaat ttgtatacgg aaacctaaag gcatcaacag gaacttgctg tgtggcgtgg 240
 124 agctaccacg ttgcaatatt ggtaagctat aaaaatgctt ccggagtaac ggaaaaaaga 300
 126 attattgatc cttcactatt ttcaagcggc cctgtaacag atacagcatg gagaaacgct 360
 128 tgcgttaaca cctcttgcg atctgcatcc gtttcctctt atgctaatac tgcaggaaat 420
 130 gtttattaca gaagtcctag taattcttac ctgtatgaca acaatctgat caataccaac 480
 132 tgtgtactga ctaaaatttc actgctttcc ggatgttctc cttcacctgc accggatgta 540
 134 tccagctgtg gattt 555
 137 <210> SEQ ID NO: 6
 138 <211> LENGTH: 185
 139 <212> TYPE: PRT
 140 <213> ORGANISM: Cryseobacterium sp. No. 9670
 142 <400> SEQUENCE: 6
 144 Leu Ala Ser Val Ile Pro Asp Val Ala Thr Leu Asn Ser Leu Phe Asn
 145 1 5 10 15
 148 Gln Ile Lys Asn Gln Ser Cys Gly Thr Ser Thr Ala Ser Ser Pro Cys
 149 20 25 30

RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/727,769A

TIME: 11:27:23

Input Set : A:\pto.vsk.txt

Output Set: N:\CRF3\02212002\I727769A.raw

```

152 Ile Thr Phe Arg Tyr Pro Val Asp Gly Cys Tyr Ala Arg Ala His Lys
153           35                      40                      45
156 Met Arg Gln Ile Leu Met Asn Asn Gly Tyr Asp Cys Glu Lys Gln Phe
157           50                      55                      60
160 Val Tyr Gly Asn Leu Lys Ala Ser Thr Gly Thr Cys Cys Val Ala Trp
161 65                      70                      75                      80
164 Ser Tyr His Val Ala Ile Leu Val Ser Tyr Lys Asn Ala Ser Gly Val
165                      85                      90                      95
168 Thr Glu Lys Arg Ile Ile Asp Pro Ser Leu Phe Ser Ser Gly Pro Val
169                      100                     105                     110
172 Thr Asp Thr Ala Trp Arg Asn Ala Cys Val Asn Thr Ser Cys Gly Ser
173                      115                     120                     125
176 Ala Ser Val Ser Ser Tyr Ala Asn Thr Ala Gly Asn Val Tyr Tyr Arg
177                      130                     135                     140
180 Ser Pro Ser Asn Ser Tyr Leu Tyr Asp Asn Asn Leu Ile Asn Thr Asn
181 145                      150                     155                     160
184 Cys Val Leu Thr Lys Phe Ser Leu Leu Ser Gly Cys Ser Pro Ser Pro
185                      165                     170                     175
188 Ala Pro Asp Val Ser Ser Cys Gly Phe
189                      180                     185
192 <210> SEQ ID NO: 7
193 <211> LENGTH: 1080
194 <212> TYPE: DNA
195 <213> ORGANISM: Cryseobacterium sp. No. 9670
197 <220> FEATURE:
198 <221> NAME/KEY: CDS
199 <222> LOCATION: (61)..(1020)
200 <223> OTHER INFORMATION:
203 <220> FEATURE:
204 <221> NAME/KEY: mat_peptide
205 <222> LOCATION: (466)..()
206 <223> OTHER INFORMATION:
209 <400> SEQUENCE: 7
210 agttaaaata accaaccaac ttaacaaaaa ctcaccatta aactacaaat tacaattatt      60
212 atg  aaa aat ctt ttt tta  tca atg atg gcc ttt  gtg acc gtc tta      105
213 Met  Lys Asn Leu Phe Leu  Ser Met Met Ala Phe  Val Thr Val Leu
214 -135                      -130                      -125
216 act  ttt aat tcc tgt gcc  gat tcc aac ggg aat  cag gaa atc aac      150
217 Thr  Phe Asn Ser Cys Ala  Asp Ser Asn Gly Asn  Gln Glu Ile Asn
218 -120                      -115                      -110
220 gga  aag gaa aaa cta agt  gta aat gat tct aag ctg aaa gat ttc gga      198
221 Gly  Lys Glu Lys Leu Ser  Val Asn Asp Ser Lys Leu Lys Asp Phe Gly
222 -105                      -100                      -95                      -90
224 aag act gta ccg gta ggg ata gac gaa gaa aac gga atg ata aag gtg      246
225 Lys Thr Val Pro Val Gly Ile Asp Glu Glu Asn Gly Met Ile Lys Val
226                      -85                      -80                      -75
228 tca ttt atg tta act gcg caa ttc tat gaa att aag ccg acc aaa gaa      294
229 Ser Phe Met Leu Thr Ala Gln Phe Tyr Glu Ile Lys Pro Thr Lys Glu
230                      -70                      -65                      -60

```

RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/727,769A

TIME: 11:27:23

Input Set : A:\pto.vsk.txt

Output Set: N:\CRF3\02212002\I727769A.raw

232	aat	gag	cag	tat	atc	gga	atg	ctt	aga	cag	gct	gtt	aag	aat	gaa	tct	342
233	Asn	Glu	Gln	Tyr	Ile	Gly	Met	Leu	Arg	Gln	Ala	Val	Lys	Asn	Glu	Ser	
234			-55					-50					-45				
236	cct	gta	cac	att	ttc	tta	aag	cct	aat	agc	aat	gaa	ata	gga	aaa	gtg	390
237	Pro	Val	His	Ile	Phe	Leu	Lys	Pro	Asn	Ser	Asn	Glu	Ile	Gly	Lys	Val	
238		-40					-35				-30						
240	gag	tct	gca	agt	ccg	gaa	gac	gta	aga	tat	ttt	aaa	acg	atc	ctg	aca	438
241	Glu	Ser	Ala	Ser	Pro	Glu	Asp	Val	Arg	Tyr	Phe	Lys	Thr	Ile	Leu	Thr	
242	-25					-20					-15				-10		
244	aaa	gaa	gta	aaa	ggg	caa	acc	aat	aaa	ttg	gcg	agt	gta	att	cct	gat	486
245	Lys	Glu	Val	Lys	Gly	Gln	Thr	Asn	Lys	Leu	Ala	Ser	Val	Ile	Pro	Asp	
246				-5				-1	1				5				
248	gta	gct	aca	tta	aat	tct	tta	ttc	aat	caa	ata	aag	aat	cag	tct	tgc	534
249	Val	Ala	Thr	Leu	Asn	Ser	Leu	Phe	Asn	Gln	Ile	Lys	Asn	Gln	Ser	Cys	
250		10					15				20						
252	ggt	acc	tct	acg	gcg	tcc	tca	cca	tgc	atc	aca	ttc	aga	tat	cct	gta	582
253	Gly	Thr	Ser	Thr	Ala	Ser	Ser	Pro	Cys	Ile	Thr	Phe	Arg	Tyr	Pro	Val	
254		25				30					35						
256	gac	gga	tgt	tat	gca	aga	gcc	cat	aag	atg	aga	caa	atc	tta	atg	aac	630
257	Asp	Gly	Cys	Tyr	Ala	Arg	Ala	His	Lys	Met	Arg	Gln	Ile	Leu	Met	Asn	
258	40					45				50					55		
260	aac	ggc	tat	gac	tgt	gaa	aaa	caa	ttt	gta	tac	gga	aac	cta	aag	gca	678
261	Asn	Gly	Tyr	Asp	Cys	Glu	Lys	Gln	Phe	Val	Tyr	Gly	Asn	Leu	Lys	Ala	
262				60					65					70			
264	tca	aca	gga	act	tgc	tgt	gtg	gcg	tgg	agc	tac	cac	gtt	gca	ata	ttg	726
265	Ser	Thr	Gly	Thr	Cys	Cys	Val	Ala	Trp	Ser	Tyr	His	Val	Ala	Ile	Leu	
266			75					80					85				
268	gta	agc	tat	aaa	aat	gct	tcc	gga	gta	acg	gaa	aaa	aga	att	att	gat	774
269	Val	Ser	Tyr	Lys	Asn	Ala	Ser	Gly	Val	Thr	Glu	Lys	Arg	Ile	Ile	Asp	
270		90					95					100					
272	cct	tca	cta	ttt	tca	agc	ggt	cct	gta	aca	gat	aca	gca	tgg	aga	aac	822
273	Pro	Ser	Leu	Phe	Ser	Ser	Gly	Pro	Val	Thr	Asp	Thr	Ala	Trp	Arg	Asn	
274		105					110					115					
276	gct	tgc	gtt	aac	acc	tct	tgc	gga	tct	gca	tcc	gtt	tcc	tct	tat	gct	870
277	Ala	Cys	Val	Asn	Thr	Ser	Cys	Gly	Ser	Ala	Ser	Val	Ser	Ser	Tyr	Ala	
278	120					125					130				135		
280	aat	act	gca	gga	aat	gtt	tat	tac	aga	agt	cct	agt	aat	tct	tac	ctg	918
281	Asn	Thr	Ala	Gly	Asn	Val	Tyr	Tyr	Arg	Ser	Pro	Ser	Asn	Ser	Tyr	Leu	
282				140					145					150			
284	tat	gac	aac	aat	ctg	atc	aat	acc	aac	tgt	gta	ctg	act	aaa	ttt	tca	966
285	Tyr	Asp	Asn	Asn	Leu	Ile	Asn	Thr	Asn	Cys	Val	Leu	Thr	Lys	Phe	Ser	
286			155					160					165				
288	ctg	ctt	tcc	gga	tgt	tct	cct	tca	cct	gca	ccg	gat	gta	tcc	agc	tgt	1014
289	Leu	Leu	Ser	Gly	Cys	Ser	Pro	Ser	Pro	Ala	Pro	Asp	Val	Ser	Ser	Cys	
290		170					175					180					
292	gga	ttt	taattaattg	ataattttac	agcacctgct	catttacaga	atcagcaggt										1070
293	Gly	Phe															
294		185															
296	gctggtatat																1080

RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/727,769A

TIME: 11:27:23

Input Set : A:\pto.vsk.txt

Output Set: N:\CRF3\02212002\I727769A.raw

```

299 <210> SEQ ID NO: 8
300 <211> LENGTH: 320
301 <212> TYPE: PRT
302 <213> ORGANISM: Cryseobacterium sp. No. 9670
304 <400> SEQUENCE: 8
306 Met  Lys Asn Leu Phe Leu  Ser Met Met Ala Phe  Val Thr Val Leu
307 -135                -130                -125
310 Thr  Phe Asn Ser Cys Ala  Asp Ser Asn Gly Asn  Gln Glu Ile Asn
311 -120                -115                -110
314 Gly  Lys Glu Lys Leu Ser  Val Asn Asp Ser Lys Leu Lys Asp Phe Gly
315 -105                -100                -95                -90
318 Lys Thr Val Pro Val Gly Ile Asp Glu Glu Asn Gly Met  Ile Lys Val
319                -85                -80                -75
322 Ser Phe Met Leu Thr Ala Gln Phe Tyr Glu Ile Lys Pro Thr Lys Glu
323                -70                -65                -60
326 Asn Glu Gln Tyr Ile Gly Met Leu Arg Gln Ala Val Lys Asn Glu Ser
327                -55                -50                -45
330 Pro Val His Ile Phe Leu Lys Pro Asn Ser Asn Glu Ile Gly Lys Val
331                -40                -35                -30
334 Glu Ser Ala Ser Pro Glu Asp Val Arg Tyr Phe Lys Thr Ile Leu Thr
335 -25                -20                -15                -10
338 Lys Glu Val Lys Gly Gln Thr Asn Lys Leu Ala Ser Val Ile Pro Asp
339                -5                -1 1 5
342 Val Ala Thr Leu Asn Ser Leu Phe Asn Gln Ile Lys Asn Gln Ser Cys
343                10                15                20
346 Gly Thr Ser Thr Ala Ser Ser Pro Cys Ile Thr Phe Arg Tyr Pro Val
347                25                30                35
350 Asp Gly Cys Tyr Ala Arg Ala His Lys Met Arg Gln Ile Leu Met Asn
351 40                45                50                55
354 Asn Gly Tyr Asp Cys Glu Lys Gln Phe Val Tyr Gly Asn Leu Lys Ala
355                60                65                70
358 Ser Thr Gly Thr Cys Cys Val Ala Trp Ser Tyr His Val Ala Ile Leu
359                75                80                85
362 Val Ser Tyr Lys Asn Ala Ser Gly Val Thr Glu Lys Arg Ile Ile Asp
363                90                95                100
366 Pro Ser Leu Phe Ser Ser Gly Pro Val Thr Asp Thr Ala Trp Arg Asn
367                105                110                115
370 Ala Cys Val Asn Thr Ser Cys Gly Ser Ala Ser Val Ser Ser Tyr Ala
371 120                125                130                135
374 Asn Thr Ala Gly Asn Val Tyr Tyr Arg Ser Pro Ser Asn Ser Tyr Leu
375                140                145                150
378 Tyr Asp Asn Asn Leu Ile Asn Thr Asn Cys Val Leu Thr Lys Phe Ser
379                155                160                165
382 Leu Leu Ser Gly Cys Ser Pro Ser Pro Ala Pro Asp Val Ser Ser Cys
383                170                175                180
386 Gly Phe
387                185
390 <210> SEQ ID NO: 9
391 <211> LENGTH: 30

```

VERIFICATION SUMMARY

DATE: 02/21/2002

PATENT APPLICATION: US/09/727,769A

TIME: 11:27:24

Input Set : A:\pto.vsk.txt

Output Set: N:\CRF3\02212002\I727769A.raw

L:83 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3

L:107 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4